International Journal of Mathematics and
Computer Applications Research (IJMCAR)
ISSN(P): 2249-6955; ISSN(E): 2249-8060
Vol. 4, Issue 4, Aug 2014, 125-136

SYNTHESIS ALGORITHMS OF FUZZY-NEURAL MODELS OF DECISION MAKING

MUHAMEDIYEVA DILNOZ TULKUNOVNA[1] & MIRZARAXMEDOVA AZIZA HUSNIDDINOVNA[2]

[1]Doctor of Technical Sciences, Chief Researcher at the Center of Development of Software and Hardware Complexes under TUIT, Uzbkeistan

[2]Candidate of Technical Sciences, Junior Researcher at the Center of Development of Software and Hardware Complexes under TUIT, Uzbkeistan

ABSTRACT

The paper considers relevant theoretical and methodological problems of development of fuzzy-neural models of decision making. Offers neuro - fuzzy algorithm of synthesis of fuzzy inference systems. Describes a two-stage adaptive synthesis algorithm of fuzzy inference systems. On the first step carried out clustering of initial fuzzy parameters in order to reduce the number of input parameters of fuzzy rules, and on the second - the synthesis of Sugeno type fuzzy models (inference rules).

KEYWORDS: Fuzzy Set, Neuro - Fuzzy Algorithm, Production Rules, Fuzzy Inference, Fuzzy Model, Knowledge Base, Expert Knowledge of the Matrix, The Algorithm, An Alternative Decision-Making

INTRODUCTION

Tasks of data mining (TAM) to identify hidden relationships between data, their classification and identifying features predictive patterns of development and analysis of the studied processes are solved in most cases, uncertainty and diverse nature of the very large volume of information. In these conditions, the results of solving such tasks are derived from the semi-structured decision making (SSDM) by using intelligent technologies of nature, including, soft computing. The greatest efficiency of solutions of these tasks is achieved when sharing their component (fuzzy logic, neural networks, and evolutionary algorithms). One of the promising areas are submitted to the research of theoretical and applied questions of synthesis of fuzzy-neural models of decision making based on the combination of theory of fuzzy sets, neural networks and evolutionary algorithms used for training and optimization of the parameters of such models.

Models and algorithms for fuzzy inference are central in decision-making, management, forecasting, classification, pattern recognition and machine learning in an uncertain fuzzy nature. An interconnected set of fuzzy inference systems implemented, which are the core of production rules like "If A, then B». These rules are formed on the basis of linguistic sentences experts.

In solving practical problems in an uncertain fuzzy nature of the information required length of construction and implementation of decision support systems can be divided into two parts: a numerical (quantitative) and linguistic (qualitative), coming from an expert. Much of the FIS uses a second type of knowledge, often presented in the form of fuzzy rule base. They show the structure of the fuzzy model of the problem as a whole and provide the basic knowledge (expert information) on the simulated system, ie the main component of "intelligence" of the problem. Therefore, the correct formation of the fuzzy rule base is essential to effectively address the problem. In order for this model was adequate to the real situation, the number of generated rules is usually FIS must be equal to the number of conditions A

rule that is number of elements in input vector system. Excessively large number of them increases the dimension of the model and, consequently, the complexity of the problem being solved. In addition, the amount of available information available, including the expert, the simulated system is often insufficient to build a more complex and adequate model. It should also take into account the existence of objective limitations on the accuracy of obtaining the source data. Therefore, their formation and evaluation in the process of building the investigated models use the principle of reasonable completeness and accuracy. This leads to the importance of systematization and classification of the initial information in order to reduce the number of reasonable rules of FIS. One of the most promising approaches to the formation of fuzzy rules and adjust the values of their parameters, especially when there are only numerical data are fuzzy neural networks (fuzzy-neural) [2]. For all the merits of their main drawback is the length of the construction of fuzzy rule base in the process of iterative learning neural networks. It is therefore appropriate to investigate the possibility of combining the methods of classification and clustering with neural network methods.

In the construction of classification and clustering procedures were reviewed and classified the various criteria [5-9].

The methods and algorithms for classification and clustering are the following: artificial neural networks, decision trees, symbolic rules, methods, nearest neighbor and k-nearest neighbor, support vector, Bayesian networks, linear regression, correlation and regression analysis, hierarchical cluster analysis methods, non-hierarchical cluster analysis methods, including algorithms for k-means and k-median, and methods of mining association rules, including the algorithm Apriori; limited exhaustive search method, evolutionary programming and genetic algorithms, a variety of methods for data visualization and many other methods.

In this classification, there are two groups of methods:

- Statistical methods based on the use of average experience, which is reflected in historical data;
- Cybernetic methods, involving many diverse mathematical approaches.

Arsenal Statistical Data mining methods classified into four groups of methods

- Descriptive analysis and a description of the source data.
- Analysis of the relations (Correlation and regression analysis, factor analysis, analysis of variance).
- Multivariate statistical analysis (Component analysis, discriminant analysis, multivariate regression analysis, canonical correlation, etc.).
- Time series analysis (Forecasting and dynamic models).
- The second area - a variety of approaches, united by the idea of computer mathematics and the use of artificial intelligence theory.

This group includes such methods:

- Artificial neural networks (Recognition, clustering, prediction);
- Evolutionary programming (Including the algorithms of the method of group accounting of arguments);
- Genetic algorithms (Optimization);

- Associative memory (Search for analogues, prototypes);
- Fuzzy logic;
- Decision trees;
- Processing expertise.

The methods designed to obtain descriptive results are iterative cluster analysis methods, including k-means algorithm, k-median, hierarchical cluster analysis methods, Kohonen self-organizing maps, cross-tabular methods of visualization, imaging, and various others. Prediction methods use the values of some variables to predict / forecast the unknown (missing) or future values of other (target) variables.

The methods aimed at obtaining predictive results, include such methods: neural networks, decision trees, linear regression, nearest neighbor, support vector, etc.

Bayesian networks (networks of trust) [5] - is an approach to classification based on a combination of Bayesian approach and graph theory. In this case, each vertex of the graph corresponds to a component of the feature vectors, arcs represent a causal relationship. Networking can be done automatically by analyzing the correlation of the components of signs. This approach does not require such strong assumptions, as the principle of maximum a posteriori probability, however, does not have the theoretical appeal, that is, in the absence of a priori data network is not built to deliver a minimum total risk. Ill-conditioning, however, is a problem for the Bayesian network, as large dimension of feature vector makes a very complex graph of relationships for the design and analysis. It also greatly increases the computational complexity. One solution to this problem is to reduce the dimension of feature vector, which leads to a deterioration of the generalization capability.

The original method of SVM was proposed by Vapnik in 1963 [6] as a method for constructing an optimal linear classifier. Although the assumption of linear separability of classes less severe than, for example, the assumption of the principle of maximum a posteriori probability, in practice it is rarely performed In 1992, proposed a way to generalize the method of support vector machines for a wide class of nonlinear problems [7]. The classical algorithm is to construct a linear separating surface (hyperplane), equidistant from the convex hulls of the classes that are based on precedents. Argued that such a separating hyperplane is optimal in terms of overall risk, with respect to any other possible hyperplanes. If such a hyperplane does not exist (not linearly separable classes), then for classification of non-linear transformation of Sound applied, projecting the original space into the space even more, possibly infinite, dimension. The linear separating surface in the Sound-induced transformation of the space is nonlinear in the original. Thus, the partially solved the problem of nonlinear classification. The algorithm of constructing the classifier is reduced to a quadratic programming problem and solved by well-known optimization methods [8]. It should be noted that the solution of quadratic programming is unique and found a global extremum.

The advantages of the algorithm k-nearest neighbors are:

- Education is to memorize the training set.
- Ease of implementation and ability to enter additional settings.
- Precedent logic of the algorithm is well understood by experts in subject areas (medicine, biometrics, law).

The disadvantage of this algorithm are

- The need to store the entire training set, which leads to inefficient memory usage.
- A large number of operations for the classification of images. As a consequence of working with large samples is much longer than traditional neural networks.

Chain-type radial basis function (RBF) has an intermediate layer of radial elements, each of which produces a Gaussian response surface. Since these functions are nonlinear, to simulate an arbitrary function is not required to take more than one intermediate layer. To simulate each function is only necessary to take a sufficient number of radial elements. It remains to solve the problem of how to combine the outputs of the hidden radial elements to obtain the output of these networks. It turns out that it is enough to take their linear combination (ie the weighted sum of Gaussian functions). RBF network has an output layer consisting of elements with linear activation functions (Haykin, 1994; Bishop, 1995). The corresponding algorithms, although faster learning algorithms MLP, to a lesser extent suitable for finding suboptimal solutions. Consequently, a model based on RBF, will run slower and require more memory than the corresponding MLP (but it is much faster than trains, and in some cases it is more important).

According to the analysis of the existing clustering methods proposed combined method of constructing fuzzy rule base using the clustering based on fuzzy relations and fuzzy neural networks.

The advantage of this method lies in its simplicity and high efficiency. In addition, it allows you to combine the numerical information provided in the form of training data from the linguistic information which is kind of rule base, by supplementing the existing data base rules that were created on the basis of numerical data.

Synthesis algorithm FIS rules and settings of their parameters is implemented in two phases.

At the first stage, the clustering (clustering) of initial input variables of the rules laid down and formed the analysis of the situation under study. Each of the formed cluster comprises a group of original input variables that are similar in some sense. In this case, each cluster is represented as a generalized condition of the relevant rules of FIS. The result is an opportunity to significantly reduce the number of generated rules FIS, corresponding to the number of clusters formed by the input variables. At this stage the rules of the preliminary, rough values of the parameters that describe the mathematical model of membership functions. In the second stage, refinement and tuning of these parameters using fuzzy neural networks and different training procedures.

To construct the fuzzy rule base using the clustering based on fuzzy relations and fuzzy neural networks, each data set was randomly divided into 10 disjoint subsets. One subset was used as the test set, the remaining nine subsets - as a set of training. The experimental results show that the proposed approach provides the classification in most cases more accurate than other methods.

Formulation of the Problem of Synthesis of the Rules of the FIS

The problem of fuzzy inference (NLV), described by a fuzzy Sugeno model [2, 10]:

$$\bigcup_{p=1}^{k_j}\left(\bigcap_{i=1}^{n} x_i = a_{i,jp} - c \;\; \text{весом} \; w_{jp}\right) \rightarrow y_j = f_j(x_1, x_2, ..., x_n). \tag{1}$$

Here: x_i, y_j - input and output variables $j = \overline{1, m}$ - number of the rule $a_{i,jp}$ - a linguistic term, which is estimated input variable x_i, in conjunction with the line-number *jp ($jp = \overline{1, k_j}$) j* - th rule $w_{jp} = [0,1]$ - rule weighting with the number of jp.The conclusion of the fuzzy rules $y_j = f(x_1, x_2, ..., x_n)$ can be described, for example, a polynomial of the form

$$y_j = b_{j,0} + b_{j,1} \cdot x_1 + b_{j,2} \cdot x_2 + ... + b_{j,n} \cdot x_n \ , \ j = \overline{1, m} \, .$$

In model (1) The input variables are estimated fuzzy terms $a_{i,jp}$ (eg, quantifiers such as OH - is very low N - low, NA - below average, C - average, BC - above average in - a tall, RH - very high), which are described own membership functions (AF). In general, the FP described by the expression:

$$\tilde{\mu}^l(x_i^j) = \frac{1}{1 + \left(\frac{x_i^j - c_l^j}{s_l^j} \right)^2} \, . \tag{2}$$

Here c_l^j, s_l^j - the parameters PT, j - number of rules $l = a_{i,jp}$ - the index term. For the specific problem under consideration formed the initial set of input variables X^0, as well as the training set as a set of pairs "input - out "with fixed (measured) values of the input $X^* = \{X_q^*\} = \{x_{1q}^*, ..., x_{nq}^*\}$ and the corresponding output parameters $Y^* = \{y_q^*\}$, $q = \overline{1, N}$ - the number of samples.

Must first reduce the set X^0 into a set X of smaller dimension and shape corresponding to the elements of rules with pre-OP values of the parameters. Then you must find by setting the parameter values c_l^j, s_l^j for which the deviation of actual values of the current rules of the conclusions of NLV (1) of the values enshrined in the training set of reference will be minimal. To solve the above problem of synthesis described below FIS developed adaptive algorithms for clustering the input variables of the model (1) and AF settings (2).

The Clustering Algorithm Input Parameters of Fuzzy Rules FIS

The main purpose of this algorithm, implemented in the first stage is to reduce the initial number of input parameters of the rules (1) and, accordingly, the number of rules, procedures using clustering and formation rules FIS with the preliminary (coarse) values of the parameters describing their PC (2). Well-known clustering algorithms K-Means and Expectation Maximization [5] impose restrictions on the geometry of clusters obtained by requiring, inter alia, the possibility of coverage for each individual cluster convex. Such a restriction imposed by these algorithms are used assumptions about the existence of centers of clusters (K-Means), or the probability density function for each cluster with the corresponding values of the expectation and variance (Expectation Maximization). Therefore, these algorithms can not adequately be divided into clusters of non-convex sets, the more sub-structures. This solves the problem described by the following algorithm for clustering a finite set of elements of arbitrary metric spaces based on partitioning the original set of equivalence classes of fuzzy relation. It allows you to group items into clusters, between which there is a sequence of

"close" to each other elements, which also corresponds to the intuitive idea of grouping Consider the clustering algorithm in the following formulation. Assume that you want to build a rule base for the FIS II with n inputs and one output. To do this, you must first create a set of pairs of training set data

$$(x_{1q}^*,\dots,x_{nq}^*,y_q^*),\ q=\overline{1,N} \tag{3}$$

Where $X_q^*=(x_{1q}^*,\dots,x_{nq}^*)$ - given the values of input parameters (conditions) of the module FIS, and y_q^* - the expected (reference) value of the output parameter (the conclusion) q - th training sample.

The essence of the clustering procedure is as follows. Suppose that: X- the metric space $d:X\to R$ - a certain metric on it; $(X_1,.X_2,\dots,X_N)\subset X$ - a sequence of elements from.

Assume that

$$\forall i\in\{1,..,N\}\exists j\in\{1,..,N\}:X_i\neq X_j \tag{4}$$

From the condition (4) that $\forall i\in\{1,\dots,N\}$ the fair:

$$\max\{d(X_i,X_k)\mid k\in\{1,..N\}\}>0 \tag{5}$$

Thus, for each index i we can define a function that describes the similarity measure the j-th element of the i-th element:

$$\xi_i:\{1,\dots,N\}\to[0,1],\ \xi_i(j):=1-\frac{d(X_i,X_j)}{\max\{d(X_i,X_k)\mid k\in\{1,\dots,N\}\}}\ .$$

For each index i, we define a function that describes the similarity measure the k-th and l-th element with respect to the i-th element:

$$\varsigma_i:\{1,\dots,N\}^2\to[0,1],$$

$$\varsigma_i(k,l):=1-\left|\xi_i(X_k)-\xi_i(X_l)\right|.$$

We now define a function describing a measure of the similarity of any two elements in the sequence $\mu(i,j)$ - with respect to all elements of the sequence:

$$\mu:\{1,\dots,N\}^2\to[0,..1],$$

$$\mu(i,j):=\min\{\varsigma_k(i,j)\mid k\in\{1,\dots,N\}\}$$

For all k and i

For k = 1, 2, ..., N recursively defined functions $\mu^{(k)}:\{1,\dots,N\}^2\to[0,1]$

$$\begin{cases} \mu^{(1)}(i,j) := \mu(i,j), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots , \\ \mu^{(k)}(i,j) := \max\{\min\{\mu^{(k-1)}(i,s), \mu^{(k-1)}(s,j)\} \mid s \in \{1,\dots,N\}\}. \end{cases}$$

$$\mu^{(k)}(i,i) = 1, \forall i,k .$$

$$\mu^{(k)}(i,j) \geq \mu^{(k-1)}(i,j), \forall k > 2 .$$

For each level $\alpha \in [0,1]$, we define a binary relation $\{X_1,.X_2,\dots,X_N\}$ on the set $R_\alpha \subset \{X_1,.X_2,\dots,X_N\}^2$ as follows:

$$(X_i, X_j) \in R_\alpha :\Leftrightarrow \mu^{(N)}(i,j) \geq \alpha .$$

Is an equivalence relation R_α which divides the set $\{X_1, X_2,\dots,X_N\}$ into disjoint equivalence classes. Two elements X_i, X_j are in the same equivalence class if and only if the value of measures of similarity (proximity) $\mu^N(i,j)$ for these elements is large

$$\mu^N(i,j) = \max(\mu^N(i,j_1),\ \mu^N(j_1,j_2),\dots\mu^N(j_r,j))$$

With the use of the proposed measures of similarity shaped fuzzy rules, which allow constructed on the basis of their strategic offensive arms to generate, for given values of input parameters, output parameters (the conclusion) with the smallest deviation of their current values from the standard set forth in the training set (3).

The clustering algorithm implemented in the following sequence.

Step 1: The division of the set $\{X_1, X_2,\dots,X_N\}$ into disjoint equivalence classes. Imagine that we know the minimum and maximum values of value each input and output information. You can identify the intervals in which there are valid values. For the input signal x_i is the interval denoted $[x_{iq}^-, x_{iq}^+]$. If the values x_{iq}^- and x_{iq}^+ are unknown, and it is possible to use the training data and choose the corretively the minimum and maximum values.

Each interval is defined in such a way to divide into regions {segments), and the value of K for each signal selecting individually, and the segments may have the same or different lengths.

To estimate the values of linguistic variables, we use the above seven-level scale of quantifiers terms. Each of these terms is a fuzzy set given by an appropriate membership function.

Using the introduced qualitative terms (classifiers) and expert knowledge, fuzzy rules represented in a table whose elements are the terms of the membership function of fuzzy rules. Using the table, and operations $\wedge$ (and - min) and $\vee$ (OR - max), it is easy to write the system of fuzzy logic equations relating the conclusions of the membership function of fuzzy inference and the input variables.

In general, each variable is the input vector $X_q^* = (x_{1q}^*, ..., x_{nq}^*)$ $q = \overline{1, N}$ has its own membership functions of fuzzy terms (OH, H, HC, C., Sun, B, S), which are used in the rules of FIS. To simplify the model we use for all the variables of the input vector only one form of membership functions.

Step 2: Construction of fuzzy rules based on the training data.

First we define the degree of membership of training data (3) to each region, selected in step 1. These degrees are expressed by the values of PT corresponding group of fuzzy sets of data.

The main advantages of the considered algorithms are:

- No a priori assumptions about the structure of the source data (Type and parameters of the probability distribution of the clusters, the centers of density, the number of clusters).
- The clear interpretation of the results of the partition of clusters: the elements are in the same cluster, when the sequence between them is close to each other elements.
- No restrictions on the geometry of clusters. Obtained with this algorithm clusters can be of any geometric shape sets, including non-convex. This distinguishes this algorithm from the known clustering algorithms (Such as modifications of the algorithms K-Means, Expectation Maximization).

Identification Algorithm and Tuning Parameters of Fuzzy Rules FIS

To identify the parameters of the conclusions of the rules (1) is proposed to use the following neuro - fuzzy algorithm:

1. Fixed values of the input and output parameters of the object state:

$$X_q^* = (x_{1q}^*, ..., x_{nq}^*),\ Y^* = \{y_q^*\}\ \ q = \overline{1, N}\ .$$

2. Determined by the values of membership functions of input parameters $\mu^q(x_i^*)$ for fixed values of the vector $X_q^* = (x_{1q}^*, ..., x_{nq}^*)$.

3. We calculate the values of membership functions of output parameters for fixed values $\mu^{y_q}\left(x_{1q}^*, x_{2q}^*, ..., x_{nq}^*\right)$ of the vector $X_q^* = (x_{1q}^*, ..., x_{nq}^*)$.

4. By training the neural network (NN) are chosen such values c_k^q, s_k^q of membership functions (2) that minimize the value of the residuals $E_t = y_q^* - y_q$, ie difference between the fixed real values of output parameters of the object (y_q^*) and values of output parameters (y_q), which are formed at the output Fuzzy NA, which approximates the rules (1). As a result, the value determined by the fuzzy output of the National Assembly, for which:

$$\mu^{y_q^*}\left(x_1^*, x_2^*, ..., x_n^*\right) = \max_{q=\overline{1,n}}\left[\mu^{y_q}\left(x_1^*, x_2^*, ..., x_n^*\right)\right]$$

For training the neuro-fuzzy network uses a system of recurrence relations, which are a modification of the error back-propagation algorithm, which for the (t +1)-th iteration of the training are as follows:

$$c_k^q(t+1) = c_k^q(t) - \eta(y_t - y_t^*)\frac{y_q\sum_{i=1}^{N}\mu^{y_i}(y_i) - \sum_{i=1}^{N} y_i\mu^{y_i}(y_i)}{\left(\sum_{i=1}^{N}\mu^{y_i}(y_i)\right)^2}\frac{1}{\mu^k(x_i^q)}\prod_{i=1}^{n}\mu^k(x_i^q)\frac{2s_k^q(x_i^q - c_k^q)^2}{((s_k^q)^2 + (x_i^q - c_k^q)^2)^2},$$

$$s_k^q(t+1) = s_k^q(t) - \eta(y_t - y_t^*)\frac{y_q\sum_{i=1}^{N}\mu^{y_i}(y_i) - \sum_{i=1}^{N} y_i\mu^{y_i}(y_i)}{\left(\sum_{i=1}^{N}\mu^{y_i}(y_i)\right)^2}\frac{1}{\mu^k(x_i^q)}\prod_{i=1}^{n}\mu^k(x_i^q)\frac{2(s_k^q)^2(x_i^q - c_k^q)}{\left((s_k^q)^2 + (x_i^q - c_k^q)^2\right)^2}.$$

The learning algorithm of neuro-fuzzy network consists of two phases. In the first phase, the model calculates the output value of the object (y), corresponding to a given network architecture. In the second phase of the residual value is calculated (E_t) and recalculated the parameters of membership functions.

Computing Experiment

A computational experiment to evaluate the effectiveness of the proposed method of classification was carried out on databases UCI machine learning. Repository UCI (UCI Machine Learning Repository) is a set of real and model of machine learning tasks, which are used by the scientific community for the empirical analysis of machine learning algorithms. Contains the actual data on applied problems in biology, medicine, physics, engineering, sociology, and other archive was established in 1987 by David Aha and colleagues graduate students at UC Irvine (School of Information & Computer Science University of California, Irvine, USA, http://www.ics.uci.edu). Since that time, it is widely used by students, faculty and researchers around the world as a major source of machine learning data sets. For the experiment were chosen six different types of real databases (Table 1).

Table 1

Database	Number of Classes	Number of Features	Number of Objects
Glass	7	9	214
Haberman	2	4	306
Iris	3	4	150
Ecology	8	7	336
Wine	3	13	178
Livers	2	6	345

Table 2 shows the comparative results of the classification accuracy of the proposed method and the methods given in [14].

Table 2

Methods of Classification					
Database	**Proposed Method**	**SVM**	**1NN**	**KNN**	**Conventional RBF Network**
Glass	87.85	71.50	72.01	72.01	69.16
Ирис	98.3	97.33	96.00	95.33	95.33
Вино	98.88	99.44	95.52	96.07	98.89

Table 3 shows the comparative results of the classification accuracy UCI databases by the proposed method and the method of SVM [16]. To construct the fuzzy rule base using the clustering based on fuzzy relations and fuzzy neural networks were variants of the calculations. Option that gives the best lower percentage of errors is accepted as a good result, and vice versa, which gives the largest percentage of error, taken as a bad result.

Table 3

Database	Proposed Method			SVM		
	Bad	**Good**	**Average**	**Bad**	**Good**	**Average**
Haberman	82.7	87.5	85.1	72.3	82.1	78.8
Liver	78.4	86	82.3	60.4	68.3	65.5
Ecoli	88.5	94.2	91.8	89.4	94.4	92.3

Table 4 shows the comparative results of the classification accuracy of the proposed method and the CBA (Classification based on associations) [15].

Table 4

Database	Proposed Method	CBA
Glass	87.85	65.28
Ирис	98.3	94.00
Вино	98.88	86.67

Experiments were performed 10 times for each data set using a 10-fold cross-validation (cross validation).

CONCLUSIONS

The proposed adaptive algorithm allows to simplify the procedures for the synthesis of fuzzy rules FIS by substantially reducing the dimension of the original variables and implement operational adjustments of fuzzy models in terms of changing environmental parameters. The results of computational experiments conducted to classify the elements of the test databases using known methods, showed a higher accuracy of the proposed two-stage adaptive algorithm for classification. The proposed approach was tested for solving the problems of forecasting using real data. The results showed high efficiency of fuzzy prediction models, synthesized by the proposed algorithm [11-13].

A promising line of research on the above issues is to develop methods and algorithms for the synthesis of the rules of the FIS using a combination of "Soft Computing" - technologies: fuzzy sets, neural networks and genetic algorithms.

REFERENCES

1. Zadeh, L.A. The concept of linguistic variable and its application to the adoption of approximate solutions. - Springer-Verlag, 1976. -165 pp.

2. Aliev R.A., Aliev R.R. The theory of intelligent systems and its application. - Baku, Publishing House of the Chashyoglu, 2001. - 720 pp.

3. Borisov A.N., Krumberg O.A., Fedorov I.P. Handling fuzzy information in decision-making systems. M.: Radio and Communications, 1989. 304 pp.

4. Pegat A. Fuzzy modeling and control. - M.: BINOM. Laboratory of Knowledge, 2009. - 798 pp. (Series "Adaptive and intelligent systems).

5. Heckerman, D. (1998) "A Tutorial on Learning with Bayesian Networks," In Jordan, M. (Ed.), Learning in Graphical Models,MIT Press.

6. Vapnik, V. and Lerner, A. J., "Generalized portrait method for pattern recognition," Automation and emote Control, vol. 24, no. 6, 1963.

7. Boser, B.; Guyon, I..; Vapnik, V. 1992. “A training algorithm for optimal margin classifiers” Fifth Annual Workshop on Computational Learning Theory. ACM Press, Pittsburgh.

8. Burges J.C. “A Tutorial on Support Vector Machines for Pattern Recognition” Knowledge Discovery and Data Mining, 2(2), 1998.

9. Vityaev E. E., Kostin V. V., Podkolodny N. A., Kolchanov N. A. Natural classifycation of nucleotide sequences. // Proc. of the Third International Conference on Bioinformatics of Genome Regulation and Structure (BGRS’2002, Novosibirsk, Russia, July 14-20, 2002), v3, ICG, Novosibirsk, 2002, pp. 197-199.

10. Shtovba S.D. Introduction to the theory of fuzzy sets and fuzzy logic. http//www.matlab.exponenta.ru.

11. Bekmuratov T.F., Muhamedieva D.T., Bobomuradov O.J. Fuzzy prediction model yields / / Journal “Problems of Informatics”. - Novosibirsk, 2010. Number 3. - P. 11-23.

12. Bekmuratov T,F., Muhamedieva D.T., Bobomuradov O.J. Model prediction of yield initial conditions. Ninth International Conference on Application of Fuzzy Systems and Soft Computing. ICAFS – 2010. . Edited by R.A. Aliev, K.W. Bonfig, M. Jamshidi, W. Pedrycz, I.B. Turksen. b – Quadrat Verlag. Prague, Czech Republic. August 26-27, 2010. – pp. 321-328.

13. Bekmuratov T, F., Muhamedieva D.T., Bobomuradov O.J. Fuzzy inference system for forecasting problems. IJUCI - International Journal of Ubiquitous Computing and Internationalization. V.3, No. 1, April 2011. ITIRC. - pp. 1-6.

14. Y.J. Oyang, S.C. Hwang, Y.Y. Ou, C.Y. Chen, Z.W. Chen, Data classification with the radial basis function network based on a novel kernel density estimation algorithm, IEEE Transactions on Neural Networks 16 (1) (2005) 225–236.

15. G. Chen, et al . A new approach to classification based on association rule mining. Decis. Support Syst. 2006; 42:674-689.

16. Y.C. Hu, Fuzzy integral-based perceptron for two-class pattern classification problems, Information Sciences 177 (2007) 1673–1686.

17. K.G. Srinivasa, K.R. Venugopal, L.M. Patnaik, A self-adaptive migration model genetic algorithm for data mining applications, Information Sciences 177 (2007) 4295–4313.